AMENDMENTS TO THE CLAIMS:

This listing of claims will replace all prior versions and listings of claims in the application:

1. (Currently Amended) A speech processing apparatus <u>which recognizes speech of a person in a car,</u> comprising:

generation means for generating a pseudo acoustic echo signal for each sample, said samples being based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

supply means for holding the current impulse response for each sample and supplying the current impulse response to said generation means;

elimination means for subtracting said pseudo acoustic echo signal from a near-end speech signal to remove an acoustic echo component and thereby generate an acoustic signal which has been echo-canceled for each sample;

update means for continually updating the impulse response for each sample by using said source signal, said acoustic echo-canceled signal and the current impulse response held by said supply means and for supplying the updated impulse response to said supply means;

decision means for checking in each frame, said frames being comprised of plurality of samples, whether or not a voice is included in the

near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled, said decision means outputting a result indicating whether said voice is included in the near-end speech signal;

storage means for storing one or more impulse responses in each frame;

control means for, in a frame for which the result of decision made by said decision means is negative, storing in said storage means the current impulse response held by said supply means and, in a frame for which the result of the decision is positive, retrieving one of the impulse responses stored in said storage means and supplying the one of the impulse responses to said supply means; ~~and~~

means for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

means for successively determining a spectrum mean for each frame based on the spectrum obtained; and

means for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source,

wherein said source signal is an output signal of a speaker of said speech processing apparatus in the car, said acoustic echo transfer path is

a path from the output signal of the speaker of said speech processing apparatus in the car to an input signal of a microphone of said speech processing apparatus in the car, said near-end speech signal is a signal of the speech of the person in the car and said additive noise of an unknown source is the car's noise with energy level of between 60 dBA and 80 dBA.

2. (Original) A speech processing apparatus as claimed in claim 1, wherein said acoustic echo-canceled signal is used for speech recognition.

3. (Cancelled).

4. (Previously Presented) A speech processing apparatus as claimed in claim 1, further comprising:

means for determining a cepstrum from the spectrum, the spectrum having the additive noise of an unknown source removed by said subtraction means;

means for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct in a lump multiplicative distortions that are dependent on

microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone, wherein said means for subtracting comprises first subtracting means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of each talker and second means for subtracting the cepstrum mean of the non-speech frame of the talker and by said first subtracting means and said second subtracting means, said subtracting means corrects in a lump multiplicative distortions that are dependent on a microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

5. (Previously Presented) A speech processing apparatus as claimed in claim 1, further comprising:

means for determining a cepstrum from the spectrum obtained; means for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone

characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

6. (Cancelled).

7. (Currently Amended) A speech processing method <u>of a speech processing apparatus which recognizes a speech of a person in a car,</u> comprising:

a generation step for generating a pseudo acoustic echo signal for each sample, said samples being based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

a supply step for holding the current impulse response for each sample and supplying the current impulse response to said generation step;

an elimination step for subtracting said pseudo acoustic echo signal from a near-end speech signal to remove an acoustic echo component and thereby generate an acoustic signal which has been echo-canceled for each sample;

an update step for continually updating the impulse response for each sample by using said source signal, said acoustic echo-canceled signal and the current impulse response held by the supply step and for supplying the updated impulse response to said supply step;

a decision step for checking in each frame, said frames being comprised of plurality of samples, whether or not a voice is included in the

near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled, said decision step outputting a result indicating whether said voice is included in the near-end speech signal;

a storage step for storing one or more impulse responses in each frame;

a control step for, in a frame for which the result of decision made by said decision step is negative, storing in said storage step the current impulse response held by the supply step and, in a frame for which the result of decision is positive, retrieving one of the impulse responses stored in said storage step and supplying it to said supply step;

a step for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

a step for successively determining a spectrum mean for each frame based on the spectrum obtained; and

a step for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source,

wherein said source signal is an output signal of a speaker of said speech processing apparatus in the car, said acoustic echo transfer path is a path from the output signal of the speaker of said speech

processing apparatus in the car to an input signal of a microphone of said speech processing apparatus in the car, said near-end speech signal is a signal of the speech of the person in the car and said additive noise of an unknown source is the car's noise with energy level of between 60 dBA and 80 dBA.

8. (Original) A speech processing method as claimed in claim 7, wherein said acoustic echo-canceled signal is used for speech recognition.

9. (Cancelled).

10. (Previously Presented) A speech processing method as claimed in claim 7, further comprising:

a step for determining a cepstrum from the spectrum removed of the additive noise;

a step for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

11. (Previously Presented) A speech processing method as claimed in claim 7, further comprising:

a step for determining a cepstrum from the spectrum obtained; a step for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

12. (Cancelled).

13. (Previously Presented) A speech processing method comprising the steps of:

applying a normalized least mean square error algorithm, controlled by a near-end talk detection algorithm based on a frame by frame basis voice activity detection algorithm, to an input signal to create an acoustic echo-cancelled signal; and

applying a continuous spectral substitution algorithm to each frame of said acoustic echo-cancelled signal to remove an unknown noise source from said acoustic echo-cancelled signal.

14. (Previously Presented) A speech processing system comprising:

means for applying a normalized least mean square error algorithm, controlled by a near-end talk detection algorithm based on a frame by frame basis voice activity detection algorithm, to an input signal to create an acoustic echo-cancelled signal; and

means for applying a continuous spectral substitution algorithm to each frame of said acoustic echo-cancelled signal to remove an unknown noise source from said acoustic echo-cancelled signal.